

## Appendix C. Source and Accuracy of Estimates

### SOURCE OF DATA

Most estimates in this report come from data obtained in March of years 1970 through 1993 in the Current Population Survey (CPS). The Bureau of the Census conducts the survey every month, although this report uses mostly March data for its estimates. Also, some estimates come from 1970 and 1980 decennial census data. The March survey uses two sets of questions, the basic CPS and the supplement.

**Basic CPS.** The basic CPS collects primarily labor force data about the civilian noninstitutional population. Interviewers ask questions concerning labor force participation about each member 15 years old and over in every sample household.

The present CPS sample was selected from the 1980 Decennial Census files with coverage in all 50 states and the District of Columbia. The sample is continually updated to account for new residential construction. It is located in 729 areas comprising 1,973 counties, independent cities, and minor civil divisions. About 60,000 occupied households are eligible for interview every month. Interviewers are unable to obtain interviews at about 2,600 of these units because the occupants are not found at home after repeated calls or are unavailable for some other reason.

Since the introduction of the CPS, the Bureau of the Census has redesigned the CPS sample several times to improve the quality and reliability of the data and to satisfy changing data needs. The most recent changes were completely implemented in July 1985.

Table C-1 summarizes changes in the CPS designs for the years for which data appear in this report.

**CPS March Supplement.** In addition to the basic CPS questions, interviewers asked supplementary questions in March about households, families, marital status, and living arrangements.

To obtain more reliable data for the Hispanic origin population, the March CPS sample was increased by about 2,500 eligible housing units, interviewed the previous November, that contained at least one sample person of Hispanic origin. In addition, the sample included

persons in the Armed Forces living off post or with their families on post.

Table C-1. Description of the March Current Population Survey

Time Period	Number of Sample Areas	Housing Units eligible <sup>1</sup>	
		Inter-viewed	Not Inter-viewed
1990 to 1993 .....	729	57,400	2,600
1989 .....	729	53,600	2,500
1986 to 1988 .....	729	57,000	2,500
1985 .....	<sup>2</sup> 629/729	57,000	2,500
1982 to 1984 .....	629	59,000	2,500
1980 to 1981 .....	629	65,500	3,000
1977 to 1979 .....	614	55,000	3,000
1973 to 1976 .....	461	46,500	2,500
1972 .....	449	45,000	2,000
1967 to 1971 .....	449	48,000	2,000
1963 to 1966 .....	357	33,500	1,500
1960 to 1962 .....	333	33,500	1,500
1957 to 1959 .....	330	33,500	1,500
1954 to 1956 .....	230	21,000	500 to 1,000
1947 to 1953 .....	68	21,000	500 to 1,000

<sup>1</sup>Excludes about 2,500 Hispanic households added from the previous November sample. (See "CPS March Supplement.")

<sup>2</sup>The CPS was redesigned following the 1980 Decennial Census of Population and Housing. During phase-in of the new design, housing units from the new and old designs were in the sample.

**CPS Estimation Procedure.** This survey's estimation procedure inflates weighted sample results to independent estimates of the civilian noninstitutional population of the United States by age, sex, race and Hispanic/non-Hispanic categories. The independent estimates were based on statistics from decennial censuses of population; statistics on births, deaths, immigration and emigration; and statistics on the size of the Armed Forces. The independent population estimates used for 1980 to present were based on updates to controls established by the 1980 Decennial Census. Data previous to 1980 were based on independent population estimates from

the most recent decennial census. For more details on the change in independent estimates, see the section entitled "Introduction of 1980 Census Population Controls" in an earlier report (Series P-60, No. 133). The estimation procedure for the March supplement included a further adjustment so that husband and wife in a household received the same weight.

The estimates in this report for 1984 and later also employ a revised survey weighting procedure for persons of Hispanic origin. In previous years, weighted sample results were inflated to independent estimates of the noninstitutional population by age, sex, and race. There was no specific control of the survey estimates for the Hispanic population. Since then, the Bureau of the Census developed independent population controls for the Hispanic population by sex and detailed age groups. Revised weighting procedures incorporate these new controls. The independent population estimates include some, but not all, undocumented immigrants.

## ACCURACY OF THE ESTIMATES

Since the CPS estimates come from a sample, they may differ from figures from a complete census using the same questionnaires, instructions, and enumerators. A sample survey estimate has two possible types of error: nonsampling and sampling. The accuracy of an estimate depends on both types of error, but the full extent of the nonsampling error is unknown. Consequently, one should be particularly careful when interpreting results based on a relatively small number of cases or on small differences between estimates. The standard errors for CPS estimates primarily indicate the magnitude of sampling error. They also partially measure the effect of some nonsampling errors in responses and enumeration but do not measure systematic biases in the data. (Bias is the average over all possible samples of the differences between the sample estimates and the desired value.)

**Nonsampling Variability.** Nonsampling errors can be attributed to several sources including the following:

- Inability to obtain information about all cases in the sample.
- Definitional difficulties.
- Differences in the interpretation of questions.
- Respondents' inability or unwillingness to provide correct information.
- Respondents' inability to recall information.
- Errors made in data collection such as in recording or coding the data.
- Errors made in processing the data.

- Errors made in estimating values for missing data.
- Failure to represent all units with the sample (undercoverage).

CPS undercoverage results from missed housing units and missed persons within sample households. Compared to the level of the 1980 Decennial Census, overall CPS undercoverage is about 7 percent. CPS undercoverage varies with age, sex, and race. Generally, undercoverage is larger for males than for females and larger for Blacks and other races combined than for Whites. As described previously, ratio estimation to independent age-sex-race-Hispanic population controls partially corrects for the bias due to undercoverage. However, biases exist in the estimates to the extent that missed persons in missed households or missed persons in interviewed households have different characteristics from those of interviewed persons in the same age-sex-race-Hispanic group. Furthermore, the independent population controls have not been adjusted for undercoverage in the 1980 census.

A common measure of survey coverage is the coverage ratio, the estimated population before ratio adjustment divided by the independent population control. Table C-2 shows CPS coverage ratios for age-sex-race groups for a recent month. The CPS coverage ratios can exhibit some variability from month to month, but these are a typical set of coverage ratios.

For additional information on nonsampling error including the possible impact on CPS data when known, refer to Statistical Policy Working Paper 3, *An Error Profile: Employment as Measured by the Current Population Survey*, Office of Federal Statistical Policy and Standards, U.S. Department of Commerce, 1978 and Technical Paper 40, *The Current Population Survey: Design and Methodology*, Bureau of the Census, U.S. Department of Commerce.

**Comparability of Data.** Data obtained from the CPS and other sources are not entirely comparable. This results from differences in interviewer training and experience and in differing survey processes. This is an example of nonsampling variability not reflected in the standard errors. Use caution when comparing results from different sources.

Caution should also be used when comparing estimates in this report, which reflect 1980 census-based population controls, with estimates for 1979 and earlier years, which reflect 1970 census-based population controls. This change in population controls had relatively little impact on summary measures such as means, medians, and percentage distributions, but did have a significant impact on levels. For example, use of 1980-based population controls results in about a 2-percent increase in the civilian noninstitutional population and in

Table C-2. CPS Coverage Ratios

Age	Non-Black		Black		All Persons		
	Male	Female	Male	Female	Male	Female	Total
0 to 14 years .....	0.948	0.960	0.913	0.930	0.943	0.955	0.949
15 years .....	0.953	0.986	0.975	1.025	0.956	0.993	0.974
16 years .....	0.877	0.997	0.886	0.963	0.879	0.991	0.934
17 years .....	0.958	0.956	0.860	0.932	0.942	0.952	0.947
18 years .....	0.950	0.958	0.931	0.692	0.947	0.916	0.931
19 years .....	0.882	0.953	0.773	0.740	0.866	0.920	0.893
20 to 24 years .....	0.889	0.918	0.645	0.820	0.856	0.904	0.881
25 to 26 years .....	0.867	0.964	0.687	0.820	0.844	0.943	0.894
27 to 29 years .....	0.919	0.941	0.700	0.834	0.892	0.926	0.909
30 to 34 years .....	0.884	0.947	0.667	0.865	0.859	0.936	0.898
35 to 39 years .....	0.892	0.936	0.693	0.928	0.871	0.935	0.903
40 to 44 years .....	0.895	0.933	0.781	0.889	0.884	0.928	0.906
45 to 49 years .....	0.933	0.955	0.842	0.938	0.925	0.953	0.939
50 to 54 years .....	0.953	0.958	0.845	0.869	0.942	0.948	0.945
55 to 59 years .....	0.918	0.905	0.797	0.906	0.906	0.905	0.905
60 to 62 years .....	0.926	0.874	0.702	0.779	0.904	0.864	0.883
63 to 64 years .....	0.851	0.960	0.814	0.944	0.848	0.959	0.906
65 to 67 years .....	0.891	0.945	0.785	0.991	0.881	0.950	0.918
68 to 69 years .....	0.876	0.986	0.741	0.810	0.864	0.970	0.922
70 to 74 years .....	0.955	1.020	0.866	0.949	0.948	1.014	0.985
75 to 99 years .....	0.983	1.019	0.713	0.861	0.962	1.006	0.990
15 years and older .....	0.911	0.951	0.752	0.877	0.893	0.942	0.919
0 years and older .....	0.919	0.953	0.802	0.891	0.905	0.945	0.926

NOTE: These coverage ratios are for May 1993.

the number of families and households. Thus, estimates of levels for data collected in 1980 and later years will differ from those for earlier years by more than what could be attributed to actual changes in the population. These differences could be disproportionately greater for certain subpopulation groups than for the total population.

Since no independent population control totals for persons of Hispanic origin were used before 1984, compare Hispanic estimates over time cautiously.

**Note When Using Small Estimates.** Summary measures (such as medians and percentage distributions) are shown only when the base is 75,000 or greater. Because of the large standard errors involved, summary measures would probably not reveal useful information when computed on a smaller base. However, estimated numbers are shown even though the relative standard errors of these numbers are larger than those for corresponding percentages. These smaller estimates permit combinations of the categories to suit data users' needs. Take care in the interpretation of small differences. For instance, even a small amount of nonsampling error can cause a borderline difference to appear significant or not, thus distorting a seemingly valid hypothesis test.

**Sampling Variability.** Sampling variability is variation that occurred by chance because a sample was surveyed rather than the entire population. Standard errors, as calculated by methods described later in "Standard Errors and Their Use," are primarily measures of sampling variability, although they may include some non-sampling error.

**Standard Errors and Their Use.** A number of approximations are required to derive, at a moderate cost, standard errors applicable to all the estimates in this report. Instead of providing an individual standard error for each estimate, generalized sets of standard errors

Table C-3. Standard Errors of Estimated Numbers

(Numbers in Thousands)

Size of Estimate	Persons	Families, Households, House- holders, and Unrelated Individuals
25 .....	12	7
50 .....	17	10
75 .....	21	12
100 .....	24	14
250 .....	38	22
500 .....	54	31
750 .....	66	38
1,000 .....	76	43
2,500 .....	120	68
5,000 .....	168	96
7,500 .....	204	117
10,000 .....	234	134
15,000 .....	282	161
25,000 .....	353	201
50,000 .....	456	260
100,000 .....	501	283
125,000 .....	460	256
150,000 .....	360	193

Note: See Table C-5 for the appropriate factor to apply to the above standard errors.

Table C-4. Standard Errors of Estimated Percentages

Base of percentage (thousands)	Estimated Percentages for Persons					
	1 or 99	2 or 98	5 or 95	10 or 90	25 or 75	50
25 .....	4.8	6.8	10.5	14.5	20.9	24.1
50 .....	3.4	4.8	7.4	10.2	14.8	17.1
75 .....	2.8	3.9	6.1	8.4	12.1	13.9
100 .....	2.4	3.4	5.3	7.2	10.4	12.1
250 .....	1.5	2.1	3.3	4.6	6.6	7.6
500 .....	1.1	1.5	2.4	3.2	4.7	5.4
750 .....	0.9	1.2	1.9	2.6	3.8	4.4
1,000 .....	0.8	1.1	1.7	2.3	3.3	3.8
2,500 .....	0.5	0.7	1.1	1.4	2.1	2.4
5,000 .....	0.3	0.5	0.7	1.0	1.5	1.7
7,500 .....	0.3	0.4	0.6	0.8	1.2	1.4
10,000 .....	0.2	0.3	0.5	0.7	1.0	1.2
15,000 .....	0.2	0.3	0.4	0.6	0.9	1.0
25,000 .....	0.2	0.2	0.3	0.5	0.7	0.8
50,000 .....	0.11	0.2	0.2	0.3	0.5	0.5
100,000 .....	0.08	0.11	0.2	0.2	0.3	0.4
125,000 .....	0.07	0.10	0.15	0.2	0.3	0.3
150,000 .....	0.06	0.09	0.14	0.2	0.3	0.3
Families, Households, Householders, and Unrelated Individuals						
25 .....	2.7	3.9	6.0	8.3	11.9	13.8
50 .....	1.9	2.7	4.2	5.8	8.4	9.7
75 .....	1.6	2.2	3.5	4.8	6.9	8.0
100 .....	1.4	1.9	3.0	4.1	6.0	6.9
250 .....	0.9	1.2	1.9	2.6	3.8	4.4
500 .....	0.6	0.9	1.3	1.8	2.7	3.1
750 .....	0.5	0.7	1.1	1.5	2.2	2.5
1,000 .....	0.4	0.6	0.9	1.3	1.9	2.2
2,500 .....	0.3	0.4	0.6	0.8	1.2	1.4
5,000 .....	0.2	0.3	0.4	0.6	0.8	1.0
7,500 .....	0.2	0.2	0.3	0.5	0.7	0.8
10,000 .....	0.14	0.2	0.3	0.4	0.6	0.7
15,000 .....	0.11	0.2	0.2	0.3	0.5	0.6
25,000 .....	0.09	0.12	0.2	0.3	0.4	0.4
50,000 .....	0.06	0.09	0.13	0.2	0.3	0.3
100,000 .....	0.04	0.06	0.09	0.13	0.2	0.2
125,000 .....	0.04	0.05	0.08	0.12	0.2	0.2
150,000 .....	0.04	0.05	0.08	0.11	0.2	0.2

Note: See Table C-5 for the appropriate factor to apply to the above standard errors.

are provided for various types of characteristics. Thus, the tables show levels of magnitude of standard errors rather than the precise standard errors.

Tables C-3 and C-4 provide standard errors of estimated numbers and estimated percentages, respectively. Table C-5 has standard error parameters for persons, families, households, householders and unrelated individuals. Tables C-6 and C-7 provide factors to apply to the standard error parameters for estimates prior to 1993.

The sample estimate and its standard error enable one to construct a confidence interval, a range that would include the average result of all possible samples with a known probability. For example, if all possible samples were surveyed under essentially the same general conditions and using the same sample design, and if an estimate and its standard error were calculated from each sample, then approximately 90 percent of the intervals from 1.645 standard errors below the estimate

to 1.645 standard errors above the estimate would include the average result of all possible samples.

A particular confidence interval may or may not contain the average estimate derived from all possible samples. However, one can say with specified confidence that the interval includes the average estimate calculated from all possible samples.

Some statements in the report may contain estimates followed by a number in parentheses. This number can be added to and subtracted from the estimate to calculate upper and lower bounds of the 90-percent confidence interval. For example, if a statement contains the phrase "grew by 1.7 percent ( $\pm 1.0$ )," the 90-percent confidence interval for the estimate, 1.7 percent, is 0.7 percent to 2.7 percent.

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The most common type of hypothesis appearing in this

report is that the population parameters are different. An example of this would be comparing the total number of family households in March 1990 to the total number of family households in March 1993.

Tests may be performed at various levels of significance, where a significance level is the probability of concluding that the characteristics are different when, in fact, they are the same. All statements of comparison in the text have passed a hypothesis test at the 0.10 level of significance or better. This means that the absolute value of the estimated difference between characteristics is greater than or equal to 1.645 times the standard error of the difference.

**Standard Errors of Estimated Numbers.** There are two ways to compute the approximate standard error,  $s_x$ , of an estimated number shown in this report. The first uses the formula

$$s_x = fs \quad (1)$$

where  $f$  is a factor from Table C-5, and  $s$  is the standard error of the estimate obtained by interpolation from Table C-3. The second method uses formula (2), from which the standard errors in Table C-3 were calculated. This formula will provide more accurate results than formula (1).

$$s_x = \sqrt{ax^2 + bx} \quad (2)$$

Here  $x$  is the size of the estimate and  $a$  and  $b$  are the parameters in Table C-5 associated with the particular type of characteristic. When calculating standard errors for numbers from cross-tabulations involving different characteristics, use the factor or set of parameters for the characteristic which will give the largest standard error.

### Illustration

Table A shows that there were 96,391,000 households in 1993. Use the appropriate parameters from Table C-5 and formula (2) to get

Estimate, $x$	96,391,000
$a$ parameter	-0.000011
$b$ parameter	1,899
Standard error	284,000
90% confidence interval	95,924,000 to 96,858,000

The standard error is calculated as

$$s_x = \sqrt{(-0.000011)(96,391,000)^2 + (1,899)(96,391,000)} = 284,000$$

The 90-percent confidence interval for the estimated number of households in 1993 is calculated as 96,391,000  $\pm$  1.645x284,000.

The alternate calculation of the standard error, using formula (1) with  $f = 1.0$  from Table C-5 and  $s = 281,000$  by interpolation from Table C-3, is

$$s_x = 1.0 \times 281,000 = 281,000$$

**Standard Errors of Estimated Percentages.** The reliability of an estimated percentage, computed using sample data for both numerator and denominator, depends on the size of the percentage and its base. Estimated percentages are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more. When the numerator and denominator of the percentage are in different categories, use the factor or parameter from Table C-5 indicated by the numerator.

The approximate standard error,  $s_{x,p}$ , of an estimated percentage can be obtained by use of the formula

$$s_{x,p} = fs \quad (3)$$

In this formula,  $f$  is the appropriate factor from Table C-5 and  $s$  is the standard error of the estimate obtained by interpolation from Table C-4.

Alternatively, formula (4) will provide more accurate results:

$$s_{x,p} = \sqrt{\frac{b}{x} p(100 - p)} \quad (4)$$

Here  $x$  is the total number of persons, families, households, or unrelated individuals which is the base of the percentage,  $p$  is the percentage ( $0 \leq p \leq 100$ ), and  $b$  is the parameter in Table C-5 associated with the characteristic in the numerator of the percentage.

### Illustration

Table A shows that 53,171,000 or 55.2 percent of the 96,391,000 households were married-couple families. Use the appropriate parameters from Table C-5 and formula (4) to get

Estimate, $p$	55.2
Base, $x$	96,391,000
$b$ parameter	1,899
Standard error	0.2
90% confidence interval	54.9 to 55.5

The standard error is calculated as

$$s_{x,p} = \sqrt{\frac{1,899}{96,391,000} (55.2)(100.0 - 55.2)} = 0.2$$

The 90-percent confidence interval for the estimated percentage of households that were married-couple families is calculated as 55.2  $\pm$  1.645x0.2.

The alternate calculation of the standard error, using formula (3) with  $f = 1.0$  from Table C-5 and  $s = 0.2$  by interpolation from Table C-4 is

$$s_{x,p} = 1.0 \times 0.2 = 0.2$$

**Standard Error of a Difference.** The standard error of the difference between two sample estimates is approximately equal to

$$s_{x-y} = \sqrt{s_x^2 + s_y^2} \quad (5)$$

where  $s_x$  and  $s_y$  are the standard errors of the estimates,  $x$  and  $y$ . The estimates can be numbers, percentages, ratios, etc. This will represent the actual standard error quite accurately for the difference between estimates of the same characteristic in two different areas, or for the difference between separate and uncorrelated characteristics in the same area. However, if there is a high positive (negative) correlation between the two characteristics, the formula will overestimate (underestimate) the true standard error.

### Illustration

Table A shows that 70.7 percent of the 96,391,000 households in 1993 were family households, whereas, 73.7 percent of the 80,776,000 households in 1980

were family households. The apparent difference between the percentage of family households in 1993 and 1980 was 3.0 percent. Using Table C-5 and formula (4) the approximate standard errors,  $s_x$  and  $s_y$ , are 0.2 and 0.2 respectively. Use formula (5) to get

	x	y	difference
Estimate	70.7	73.7	3.0
Standard error	0.2	0.2	0.3
90% confidence interval	70.4 to 71.0	73.4 to 74.0	2.5 to 3.5

The approximate standard error of the difference is calculated as

$$s_{x-y} = \sqrt{0.2^2 + 0.2^2} = 0.3$$

The 90-percent confidence interval for the estimated difference between the percentage of households which were family households in 1993 and 1980 is calculated as  $3.0 \pm 1.645 \times 0.3$ . Because this interval does not

**Table C-5. Parameters and Factors for Persons, Families, Households, Householders, and Unrelated Individuals: March 1993**

Characteristics	a	b	f
<b>NONINCOME CHARACTERISTICS</b>			
<b>Persons</b>			
Total or White			
Some Household Members	-0.000026	4,785	0.9
All Household Members	-0.000033	5,815	1.0
Black			
Some Household Members	-0.000283	6,864	1.1
All Household Members	-0.000417	10,121	1.3
Asian or Pacific Islanders and Other Races			
Some Household Members	-0.000719	6,864	1.1
All Household Members	-0.001060	10,121	1.3
Hispanic Origin			
Some Household Members	-0.000567	6,864	1.1
All Household Members	-0.000836	10,121	1.3
<b>Families, Households, Householders, and Unrelated Individuals</b>			
Total or White	-0.000011	1,899	1.0
Black	-0.000071	1,716	1.0
Asian or Pacific Islanders and Other Races	-0.000180	1,716	1.0
Hispanic	-0.000142	1,716	1.0
<b>INCOME CHARACTERISTICS</b>			
<b>Families, Households, Householders, and Unrelated Individuals</b>			
Total or White	-0.000012	2,058	1.0
Black	-0.000109	2,243	1.1
Asian or Pacific Islanders and Other Races	-0.000322	2,243	1.1
Hispanic	-0.000175	2,243	1.1
<b>BELOW POVERTY LEVEL</b>			
<b>Families, Households, Householders, and Unrelated Individuals</b>			
Total or White	-0.000093	2,243	1.1
Black	-0.000093	2,243	1.1
Asian or Pacific Islanders and Other Races	-0.000093	2,243	1.1
Hispanic	-0.000093	2,243	1.1

Notes: To obtain parameters prior to 1993, multiply by the appropriate factors in Tables C-6 and C-7. For regional estimates multiply the a and b parameters by 0.74, 0.98, 1.04, and 1.06 for Northeast, Midwest, South, and West, respectively. The a and b parameters should be multiplied by 1.5 for nonmetropolitan residence categories and 1.9 for farm categories.

contain zero, we can conclude with 90-percent confidence that the percentage of households which are family households has decreased between 1980 and 1993.

**Standard Error of a Median.** The sampling variability of an estimated median depends on the form of the distribution and the size of the base. One can approximate the reliability of an estimated median by determining a confidence interval about it. (See the section "Standard Errors and Their Use" for a general discussion of confidence intervals.)

Estimate the 68-percent confidence limits of a median based on sample data using the following procedure.

1. Determine, using formula (4), the standard error of the estimate of 50 percent from the distribution.
2. Add to and subtract from 50 percent the standard error determined in step 1.
3. Using the distribution of the characteristic, determine upper and lower limits of the 68-percent confidence interval by calculating values corresponding to the two points established in step 2. Use the following formula to calculate the upper and lower limits.

$$X_{pN} = \frac{pN - N_1}{N_2 - N_1} (A_2 - A_1) + A_1 \quad (6)$$

where  $X_{pN}$  = estimated upper and lower bounds for the confidence interval ( $0 \leq p \leq 1$ ). For purposes of calculating the confidence interval,  $p$  takes on the

values determined in step 2. Note that  $X_{pN}$  estimates the median when  $p = 0.50$ .

- $N$  = for distribution of numbers: the total number of units (persons, households, etc.) for the characteristic in the distribution.  
 = for distribution of percentages: the value 1.0.
- $p$  = the values obtained in step 2.
- $A_1, A_2$  = the lower and upper bounds, respectively, of the interval containing  $X_{pN}$ .
- $N_1, N_2$  = for distribution of numbers: the estimated number of units (persons, households, etc.) with values of the characteristic greater than or equal to  $A_1$  and  $A_2$ , respectively.  
 = for distribution of percentages: the estimated percentage of units (persons, households, etc.) having values of the characteristic greater than or equal to  $A_1$  and  $A_2$ , respectively.

4. Divide the difference between the two points determined in step 3 by two to obtain the standard error of the median.

Use of the above procedure could result in standard errors which differ from those given in the detailed tables. The reasons for this discrepancy are the use of a more detailed distribution than that given in the tables in determining the published standard errors, and the rounding of the numbers to thousands in the published tables.

## Illustration

Table D shows that the estimated median age of all householders was 45.9 in 1993. Table D also shows that the size, or base, of the distribution from which this median was determined was 96,391,000 householders.

1. Using formula (4) with  $b = 1,899$  from Table C-5, the standard error of 50 percent on a base of 96,391,000 is about 0.2 percent.
2. To obtain a 68-percent confidence interval on the estimated median, add to and subtract from 50 percent the standard error found in step 1. This yields percentage limits of 49.8 and 50.2.
3. Table D also shows that 49,909,000 householders were 45 years of age or older and 33,333,000 were 55 or older. Using formula (6), the lower limit for the confidence interval of the median is:

$$\frac{0.502 \times 96,391,000 - 49,909,000}{33,333,000 - 49,909,000} (55 - 45) + 45 = 45.9$$

Similarly, the upper limit can be computed as

Table C-6. **Factors to Approximate Parameters Prior to 1993: Total, White, Black, and Asian or Pacific Islanders**

Time Period	Factor
1990-1992 .....	1.00
1989 .....	1.11
1982-1988 .....	0.94
1977-1981 .....	0.84
1976 or earlier .....	0.82

Note: Apply the appropriate factor to the parameters in Table C-5.

Table C-7. **Factors to Approximate Parameters Prior to 1993: Hispanic Origin**

Time Period	Families	Persons	
		Some	All
1990-1992 .....	1.00	1.00	1.00
1989 .....	1.30	1.30	1.30
1985-1988 .....	0.94	0.94	0.94
1982-1984 .....	1.06	0.83	1.13
1977-1981 .....	0.95	0.74	1.01
1976 or earlier .....	0.92	0.72	0.98

Note: Apply the appropriate factor to the parameters in Table C-5.

$$\frac{0.498 \times 96,391,000 - 49,909,000}{33,333,000 - 49,909,000} (55 - 45) + 45 = 46.2$$

Thus, a 68-percent confidence interval for the median age of all householders is from 45.9 to 46.2.

4. Therefore, the standard error of the median is

$$S_{\text{median}} = \frac{46.2 - 45.9}{2} = 0.2$$

**Standard Error of a Ratio.** Certain estimates may be calculated as the ratio of two numbers. The standard error of a ratio,  $x/y$ , may be computed using

$$s_{x/y} = \frac{x}{y} \sqrt{\left[\frac{s_x}{x}\right]^2 + \left[\frac{s_y}{y}\right]^2 - 2r \frac{s_x s_y}{xy}} \tag{7}$$

The standard error of the numerator,  $s_x$ , and that of the denominator,  $s_y$ , may be calculated using formula (2). Alternatively, use formula (1) and Table C-5. In formula (7),  $r$  represents the correlation between the numerator and the denominator of the estimate.

For one type of ratio, the denominator is a count of families or households and the numerator is a count of persons in those families or households with a certain characteristic. If there is at least one person with the characteristic in every family or household, use 0.7 as an estimate of  $r$ . An example of this type is the mean number of children per family with children.

For all other types of ratios,  $r$  is assumed to be zero. If  $r$  is actually positive (negative), then this procedure will provide an overestimate (underestimate) of the

standard error of the ratio. Examples of this type are the mean number of children per family and the poverty rate for families.

NOTE: For estimates expressed as the ratio of  $x$  per 100 $y$  or  $x$  per 1,000 $y$ , multiply formula (7) by .100 or 1,000, respectively, to obtain the standard error.

**Illustration**

Table 16 shows the total population in nonfamily households,  $x$ , was 34,959,000 and the total number of nonfamily households,  $y$ , was 28,247,000 in 1993. The average number of persons per nonfamily household is 1.24. Use formula (2) to compute the standard error of  $x$  and  $y$ .

	x	y	ratio
Estimate	34,959,000	28,247,000	1.24
Standard error	404,000	212,000	0.01
90% confidence interval	34,294,420 to 35,623,580	27,898,260 to 28,595,740	1.22 to 1.26

Using formula (7), the standard error is calculated as

$$s_{x/y} = \frac{34,959,000}{28,247,000} \sqrt{\left[\frac{404,000}{34,959,000}\right]^2 + \left[\frac{212,000}{28,247,000}\right]^2 - 2 \times 0.7 \times \frac{404,000 \times 212,000}{34,959,000 \times 28,247,000}} = 0.01$$

The 90-percent confidence interval for the average number of persons per nonfamily household is calculated as  $1.24 \pm 1.645 \times 0.01$ .